

THE USE OF ARTIFICIAL NEURONAL NETWORKS OF THE *RBF* TYPE FOR PREDICTION OF YIELD OF CHOSEN CEREAL PLANTS

Summary

*Appearing recently methods, having guilds of artificial intelligence, permit on building of simulating models which realize assigned tasks on the basis of patterns taken directly with nature observation [1]. The processing techniques based on artificial neural networks create a special group, being in fact a computer simulators of brain work [3]. With the help of neuronal models it is possible to predict the expected crops yield on the basis of empirical data regarding crop yields in last summers. This work proposes utilization of prediction methods, which represent chosen topologies of neuronal nets among others, the *RBF* (**Radial Basis Functions**) neural network peculiarly.*

WYKORZYSTANIE SZTUCZNYCH SIECI NEURONOWYCH TYPU *RBF* DO PREDYKCJI PŁONU WYBRANYCH ROŚLIN ZBOŻOWYCH

Streszczenie

*Pojawiające się ostatnio metody, mające cechy sztucznej inteligencji, pozwalają na budowę modeli symulacyjnych, które realizują postawione zadania w oparciu o wzorce zaczerpnięte bezpośrednio z obserwacji przyrody [1]. Szczególną grupę stanowią techniki przetwarzania oparte na sztucznych sieciach neuronowych, będące w istocie komputerowymi symulatorami pracy mózgu [3]. Za pomocą modeli neuronowych można m.in. dokonać predykcji wielkości plonów płodów rolnych w oparciu o posiadane empiryczne dane, dotyczące zbiorów w latach ubiegłych. W pracy proponuje się wykorzystanie technik predykcyjnych, jakie m.in. reprezentują wybrane topologie sieci neuronowych, w szczególności sieci neuronowe typu *RBF* (**Radial Basis Functions**).*

Wstęp

Sztuczne sieci neuronowe w ostatnich latach stały się obiektem intensywnych badań prowadzonych w wielu obszarach naukowych. Wysoka wydajność i szybkość działania modeli neuronowych implikuje szerokie widmo zastosowań w różnych dziedzinach życia. Działania uczonych, wykorzystujących w swoich badaniach modelowanie neuronowe, skupiają się m.in. na przyspieszeniu obliczeń czasu rzeczywistego. Funkcjonalność taką dają im sztuczne sieci neuronowe, charakteryzujące się wielowątkowym oraz rozproszonym procesem przetwarzania informacji. Dzięki tej właściwości można wykorzystać modelowanie neuronowe m.in. w predykcji, klasyfikacji, kojarzeniu, analizie, filtracji oraz szeroko rozumianej optymalizacji. Coraz powszechniej wykorzystywanymi w praktyce topologiami sieci neuronowych, są sieci typu *RBF* (**Radial Basis Function**).

Sieci *RBF* posiadają kilka istotnych zalet w porównaniu z sieciami najczęściej używanymi, jak np. sieci typu *MLP* (**MultiLayer Perceptrons**) [3]. Po pierwsze mogą one modelować dowolną funkcję nieliniową za pomocą pojedynczej warstwy ukrytej. Eliminuje to, już na etapie projektowania sieci neuronowej, konieczność podejmowania trudnej zazwyczaj decyzji dotyczącej topologii sieci, a w szczególności określenie liczby warstw ukrytych. Po drugie, prosta transformacja liniowa dokonywana w warstwie wyjściowej sieci typu *RBF* może być w całości zoptymalizowana przy użyciu tradycyjnych technik modelowania liniowego, które są szybkie i przy stosowaniu których nie pojawiają się takie problemy jak np. minima lokalne. Są one jak wiadomo szczególnym utrudnieniem występującym

w trakcie uczeniu sieci typu *MLP*. Z tego powodu sieci *RBF* mogą być uczone w bardzo krótkim czasie (różnica w szybkości uczenia dotyczy rzędów wielkości). Cechą różniącą działanie sieci *RBF* od sieci *MLP* jest przede wszystkim inne podejście do modelowania przestrzeni wejść i wag. Model uzyskiwany w przypadku sieci *RBF* można określić jako "skupieniowy" (lokalny) natomiast model uzyskiwany w przypadku sieci *MLP* można identyfikować jako "płaszczyznowy" (globalny). Ta odmienność modeli ma swoje praktyczne konsekwencje. Należy zauważyć, że bardziej wyrafinowane kształty modelowanych funkcji i na ogół wyższe oczekiwania odnośnie dokładności generowanych odpowiedzi wymagają zastosowania w sieciach *RBF* większej liczby neuronów w stosunku do sieci typu *MLP*. Wskutek tego, modele oparte na sieciach *RBF* mają skłonność do wolniejszego działania i wymagają większych obszarów pamięci niż odpowiadające im sieci *MLP*. Należy jednak stwierdzić, że sieci *RBF* zadowolają się krótszym czasem uczenia, co w niektórych zastosowaniach może być bardzo istotnym atrybutem.

Charakterystyka sieci typu *RBF*

Istotą topologii sieci neuronowej typu *RBF* jest koncepcja zastosowania nieliniowego przekształcenia przestrzeni danych wejściowych poprzez wprowadzenie warstwy neuronów ukrytych realizujących funkcję zmieniającą się radianie wokół wybranego *a priori* centrum c_i . Sieć o radialnych funkcjach bazowych zaproponowana została przez *Broomhead'a* i *Lowe'a* (1988) oraz *Moody'a* i *Darkina* (1989). Sieci radialne *RBF* reprezentują odmienny, w sto-

sunku do sieci sigmoidalnych typu *MLP*, sposób odwzorowania zbioru wejściowego w zbiór wyjściowy. Odwzorowanie to polega na dopasowaniu funkcji aproksymującej wielu zmiennych do wartości żądanych, czyli rozpięciu nad zbiorem uczącym wielowymiarowej hiperpowierzchni dopasowującej się do wartości żądanych. Sieci *RBF* wymagają zazwyczaj użycia większej liczby neuronów niż sieci typu perceptron, jednak uczenie ich przebiega o wiele szybciej [5].

Podstawą teoretyczną działania sieci *RBF* jest twierdzenie Covera o separowalności wzorców. Mówi ono, że złożony problem klasyfikacyjny rzutowany nieliniowo na przestrzeń wielowymiarową, może być rozdzielony za pomocą separatora liniowego z większym prawdopodobieństwem niż w wypadku rzutowania na przestrzeń o mniejszej liczbie wymiarów [2],[3]. Oznacza to, że złożony problem klasyfikacyjny odwzorowany nieliniowo w wysoko wymiarowej przestrzeni ma większe szanse (większe prawdopodobieństwo) być liniowo separowalny niż w oryginalnej przestrzeni, charakteryzującej się niższym wymiarem.

Sieć neuronowa typu *RBF* funkcjonuje analogicznie do działania wielowymiarowej interpolacji, odwzorowując wektory wejściowych x_i (dla $i = 1, 2, 3, \dots, p$) z przestrzeni n -wymiarowej w zbiór liczb rzeczywistych d_i (dla $i = 1, 2, 3, \dots, p$). Zapis ten jest tożsamy z problemem określenia takiej funkcji interpolacyjnej $f(x)$, dla której spełniony będzie warunek:

$$f(x_i) = d_i \quad (1)$$

przy jednoczesnym założeniu reguły aktywacji właściwej dla sieci typu *RBF* postaci:

$$f(x) = \sum_{i=1}^p w_i \cdot \varphi(\|x - c_i\|) \quad (2)$$

gdzie:

- w_i jest wektorem wag sieci, podlegającym modyfikacji w procesie uczenia,
- zapis $\|\dots\|$ oznacza normę (najczęściej euklidesową),
- wartości wektorów c_i (dla $i = 1, 2, 3, \dots, p$) stanowią centra funkcji radialnych,
- x_i reprezentują węzły interpolacji zakodowane w postaci wektorów wejściowych sieci reprezentujące dane (np. empiryczne).

Neurony ukryte realizują więc funkcje zmieniające się radialnie wokół wybranego centrum c . Funkcje takie oznaczane ogólnie w postaci $\varphi(\|x - c\|)$ nazywają się radialnymi funkcjami bazowymi. Rola neuronu ukrytego sprowadza się zatem do odwzorowania radialnego przestrzeni wokół jednego zadanego punktu (lub grupy takich punktów).

Sieci typu *RBF* stanowią w istocie rozwinięcie oraz uogólnienie koncepcji sieci sigmoidalnych. Neuron sigmoidalny reprezentuje w przestrzeni wielowymiarowej hiperpłaszczyznę separującą w sposób globalny tę przestrzeń na dwie klasy. Neuron radialny natomiast generuje hipersferę, dokonującą podziału kołowego (lokalnej separacji radialnej) wokół punktu centralnego [4]. Pod tym względem stanowi więc on naturalne uzupełnienie neuronu sigmoidalnego, umożliwiające w przypadku wystąpienia symetrii kołowej w zbiorze danych (wejściowych) wydajne zmniejszenie

liczby neuronów, potrzebnych do realizacji zadania klasyfikacyjnego. W rozwiązaniach praktycznych sieć *RBF* jest strukturą dwuwarstwową (wg. notacji bez uwzględnienia zdeterminowanej warstwy wejściowej), w której jedynie warstwa ukryta reprezentuje odwzorowanie nieliniowe, realizowane przez neurony o radialnej funkcji bazowej. Uczenie sieci typu *RBF* jest procesem nieco bardziej złożonym, niż w przypadku uczenia perceptronu. Proces ten przebiega zazwyczaj w trzech rozłącznych fazach: wyznaczenie centrów, określenie odchyleń oraz optymalizacja liniowa warstwy wyjściowej.

Symulacja modelu neuronowego

Programowe symulatory sieci neuronowych w praktyce same dostrajają się tworząc poszukiwane modele, ponieważ automatycznie uczą się na podanych przykładach. Odbywa się to w taki sposób, że użytkownik sieci gromadzi reprezentatywne dane pokazujące, jak manifestują się badane zależności. Następnie uruchamia mechanizm (algorytm) uczenia, który ma na celu automatyczne wytworzenie w pamięci sieci właściwej struktury danych, zakodowanych w specyficzny sposób w topologii sieci. Opierając się na tej samodzielnie wytworzonej strukturze danych („zapamiętanych” w postaci tzw. współczynników wagowych, dostrojonych w trakcie procesu uczenia) sieć realizuje następnie pożądane funkcje związane z eksploatacją wygenerowanego modelu neuronowego. Użytkownik potrzebuje jednak pewnej, w głównej mierze empirycznej, wiedzy dotyczącej sposobu wyboru i przygotowania danych stanowiących przykłady uczące. Musi także dokonać wyboru właściwego rodzaju sieci neuronowej, określić typ algorytmu uczącego oraz na koniec zadać sobie trud interpretacji otrzymanych rezultatów.

Do wytworzenia, testowania oraz weryfikacji modelu neuronowego typu *RBF*, służącego do predykcji plonu wybranych roślin zbożowych, wykorzystano platformę *Statistica Neural Networks v. 4.0*. Pakiet *SNN v. 4.0* jest przyjaznym użytkownikowi produktem zaawansowanym technologicznie, charakteryzującym się wyjątkową stabilnością oraz wysoką sprawnością działania. Oferuje on przy tym wiele unikalnych narzędzi analitycznych, pozwalających na kompleksowe wytwarzanie oraz testowanie złożonych topologii neuronowych.

Znalezienie odpowiedniego modelu neuronowego adekwatnego do zadanego problemu, okazało się żmudne i czasochłonne. Seria eksperymentów symulacyjnych pozwoliła na wygenerowanie, a następnie wyselekcjonowanie oraz przetestowanie zbioru sieci typu *RBF* o najlepszych parametrach realizowanych zarówno na zbiorze walidacyjnym jak również testowym. Sukcesem okazały się sztuczne sieci neuronowe typu *RBF* o strukturach różnicowanych adekwatnie dla poszczególnych roślin. Potrafiły one w sposób satysfakcjonujący dokonać przewidywania wielkości plonów trzech wybranych roślin zbożowych: jęczmienia, pszenicy i żyta. Wygenerowane modele dokonywały predykcji plonów w oparciu o wielkości zadane na wejściu, stanowiące podstawowe dane warunkujące rozwój tych roślin, takich jak: temperatura, opady oraz usłonecznienie. Historyczne dane dotyczące plonów w latach 1995 – 2003 pozyskano z internetowej witryny GUS. Przykładowe dane uczące zapamiętane w strukturze edytora danych pakietu *SNN v. 4.0* dla przypadku żyta przedstawia rys. 1.

	TEMPER	OPADY	USIONEC	MIESIAC	PLON
94	7,3	93	68	10	24,6
95	4,3	39	46	11	24,6
96	-4,9	16	47	12	24,6
97	-2,3	35	42	1	21,4
98	-4,2	10	48	2	21,4
99	2,3	18	145	3	21,4
100	7,2	32	187	4	21,4
101	15,1	60	260	5	21,4
102	17,9	39	299	6	21,4
103	19,3	94	112	7	21,4
104	18,8	35	277	8	21,4
105	13,9	38	200	9	21,4
106	5,6	57	94	10	21,4
107	5,4	26	58	11	21,4
108	1,3	44	48	12	21,4

Rys. 1. Edytor zbioru danych dla żyta
Fig. 1. Editor of data for rye

Jako cechy reprezentatywne przyjęto 4 zmienne wejściowe: średnią temperaturę miesięczną, średnie opady w miesiącu, średnie miesięczne nasłonecznienie oraz miesiąc zakodowany w zakresie liczb 1-12. Wyjście sieci stanowił oczywiście poszukiwany plon rośliny. W procesie uczenia sieci wykorzystano 108 przypadków podzielonych losowo na trzy zbiory: uczący (57 przypadków oznaczonych w edytorze kolorem czarnym), walidacyjny (36 przypadków oznaczonych w edytorze kolorem czerwonym) oraz testowy (15 przypadków oznaczonych w edytorze kolorem niebieskim).

Proces uczenia sieci *RBF* realizowano w trzech etapach:
 - wyznaczenie centrów: zrealizowano metodą k-średnich,
 - określenie odchyłeń: zrealizowano metodą k-sasiadów,
 - optymalizację warstwy wyjściowej: zrealizowano metodą pseudoinwersji.

Otrzymano trzy zbiory sieci neuronowych z których wyselekcjonowano następujące topologie sieci *RBF*, charakteryzujące się najlepszą jakością (popołniające najmniejszy błąd predykcji):

- dla żyta jest to model *RBF* o strukturze (4-2-1),
- dla pszenicy jest to model *RBF* o strukturze (4-10-1),
- dla jęczmienia jest to model *RBF* o strukturze (4-8-1).

Przykładowe statystyki regresyjne dla sieci modelującej plon żyta przedstawiono na rys. 2.

	Uc. PLON	Wa. PLON	Te. PLON
Średnia	23,3434	22,84315	22,94335
Odch. std.	9,612241	9,762344	9,128069
Średni błąd	0,011179	-0,1389	-0,1238
Odch. błąd	2,011181	1,744864	1,875685
Śr. bł. bez	1,445855	1,322413	1,413262
Iloraz odch	0,2092313	0,1787341	0,2054854
Korelacja	0,9778679	0,9839028	0,9786935

Rys. 2. Statystyki regresyjne dla żyta
Fig. 2. Regression statistics for rye

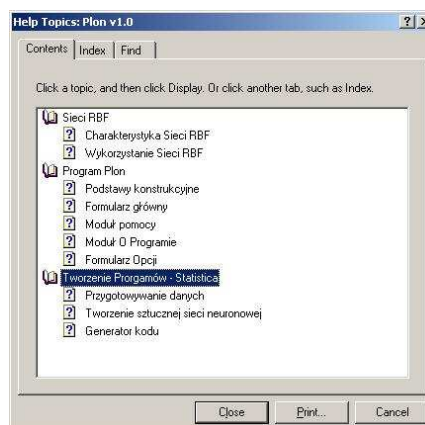
Najważniejsze statystyki umieszczono u dołu tabeli. Iloraz odchyłeń standardowych dla błędów sieci oraz danych, będący głównym wskaźnikiem jakości zbudowanego neuronowego modelu regresyjnego, jest na poziomie 0,2. Świadczy to o dobrej realizacji regresji przez wytworzoną sieć neuronową. Można zauważyć, że wysoka korelacja (standardowy współczynnik korelacji *R Pearsona* dla wartości zadanej oraz wartości uzyskanej na wyjściu sieci) na poziomie 0,97 jest zbliżona zarówno dla zbioru uczącego, walidacyjnego jak i testowego. Wysoka korelacja dla zbioru walidacyjnego oraz testowego świadczy o zdolnościach do generalizacji (predykcji) wykreowanego modelu neuronowego (sieć nie wykazuje tendencji do przeuczenia, tzn. nie jest nauczona „na pamięć”).

Informatyczny system edukacyjny

W celu ułatwienia eksploatacji wygenerowanych sieci neuronowych jako instrumentu predykcyjnego (m.in. w celu uniezależnienia się od platformy *SNN v.4.0*) została wytworzona aplikacja, która może mieć charakter zarówno użytkownikowy jak również edukacyjny. Wykorzystując mechanizmy optymalizacyjne zaimplementowane w programie *SNN v. 4.0* dokonano, wykorzystując techniki oparte na algorytmach genetycznych, weryfikacji liczby zmiennych wejściowych. Następnie korzystając z generatora kodu oferowanego komercyjnie w postaci modułu pakietu *Statistica v.6.0* oraz wykorzystując środowisko programistyczne *Visual C++*, wytworzono system informatyczny o nazwie „Plon v.1.0”.



Rys. 3. Interfejs programu "Plon v.1.0"
Fig. 3. Interface of "Plon v.1.0" program



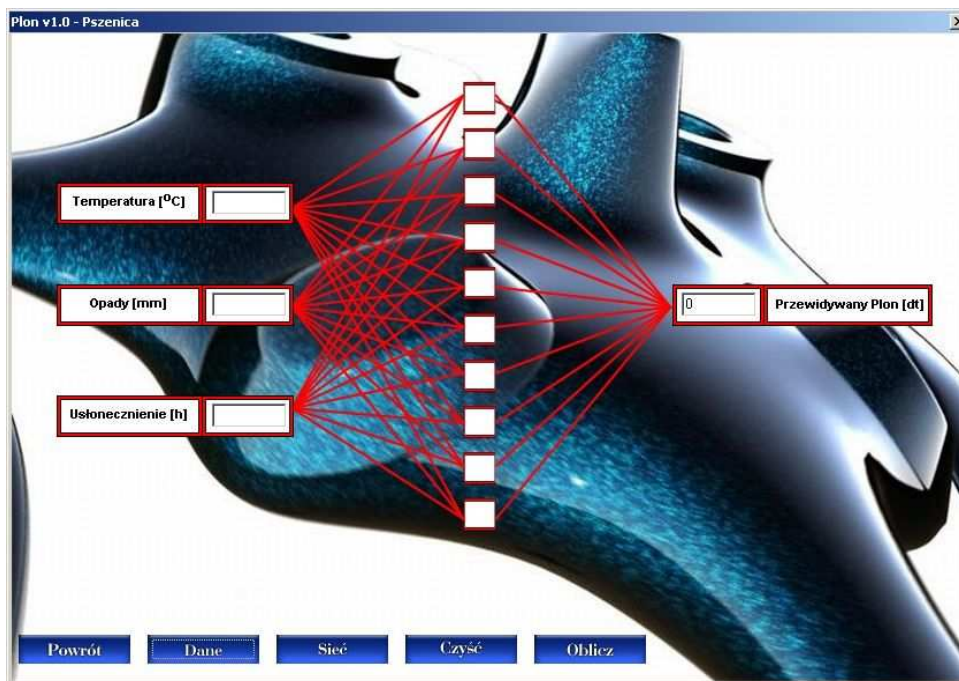
Rys. 4. Moduł pomocy programu "Plon v.1.0"
Fig. 4. Module Help of program "Plon v.1.0"

Struktura programu „Plon v.1.0” składa się z formularzy odpowiednio rozmieszczonych i dostępnych poprzez formularz główny. Jest on zaopatrzony w przyciski umożliwiające szybki dostęp do wybranych opcji programu.

Przykładowo moduł „Pomoc” posiada trzy poziomy. Pierwszy charakteryzuje sieci RBF, użyte do wytworzenia oprogramowania, ich właściwości, a także możliwe implementacje. Drugi poziom opisuje działanie programu i moż-

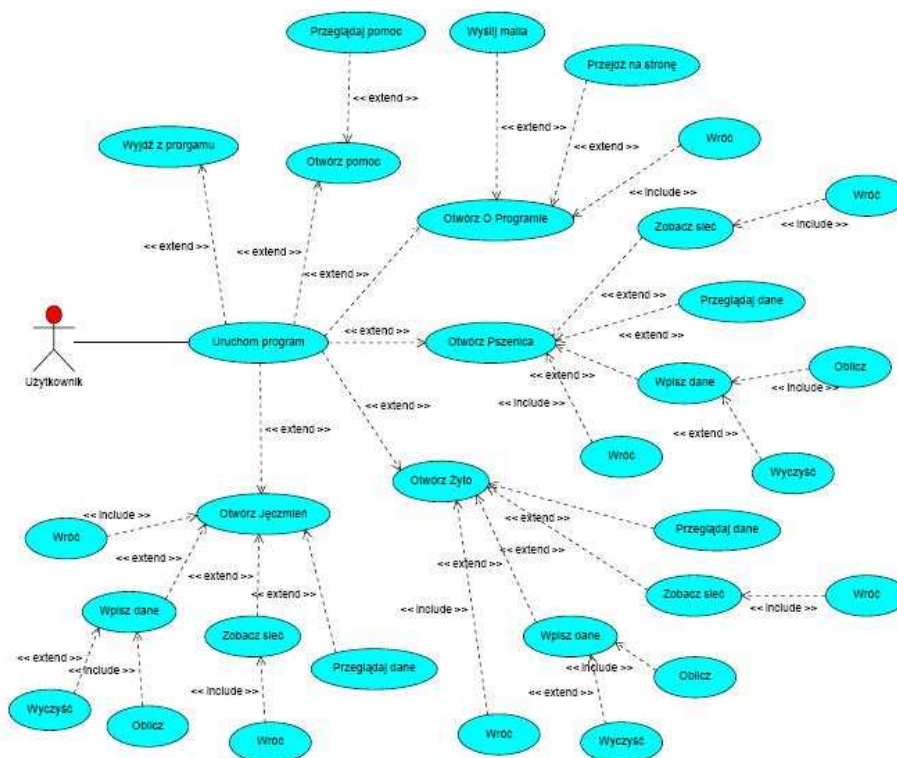
liwości aplikacji. Trzeci poziom, przedstawia uproszczony sposób utworzenia aplikacji, na podstawie posiadanych danych i programu „Statistica”.

Poniżej przedstawiono przykładowo okno interaktywnego modułu „Pszenica” służącego do neuronowej predykcji wielkości plonu pszenicy, w oparciu o wprowadzone dane.



Rys. 5. Moduł „Pszenica” programu "Plon v.1.0"
Fig. 5. Module „Pszenica” of program "Plon v.1.0"

Diagram przypadków użycia programu "Plon v.1.0" przedstawiono na rys. 5.



Rys. 6. Diagram przypadków użycia programu "Plon v.1.0"
Fig. 6. Diagram of use cases for program "Plon v.1.0"

W celu weryfikacji poprawności działania systemu dokonano serii testów wykorzystując komputery o zróżnicowanych parametrach oraz pracujących w oparciu o różne systemy operacyjne. Pomyślnie zakończone testy wykazały dużą szybkość działania, stabilność i niezawodność wytworzonej aplikacji.

Uwagi końcowe

Sztuczne sieci neuronowe stanowią intensywnie rozwijającą się dziedzinę wiedzy, stosowaną w wielu obszarach nauki oraz praktyki. Mają właściwości pożądane w wielu zastosowaniach praktycznych, stanowiąc jednocześnie uniwersalny układ aproksymacyjny, odwzorowujący wielowymiarowe zbiory danych. Mają zdolność uczenia się i adaptacji do zmieniających się warunków środowiskowych. Posiadają też umiejętność uogólniania nabytej wiedzy, stanowiąc pod tym względem system sztucznej inteligencji.

Na podstawie pozyskanych danych empirycznych oraz wykorzystując technikę przetwarzania danych, właściwą dla sieci typu *RBF*, opracowano system informatyczny "Plon v.1.0" dokonujący predykcji średniego plonu pszenicy, żyta oraz jęczmienia. Poprzez rozbudowaną strukturę plików pomocy zawartej w programie umożliwiono również zapoznanie się z praktycznymi możliwościami tego rodzaju sieci, oraz wykorzystaniu ich do predykcji innego rodzaju danych.

Przeprowadzone badania symulacyjne pozwalają na wyciągnięcie następujących wniosków:

- Sieci neuronowe typu *RBF* reprezentują właściwą topologię neuronowego modelu jako instrumentu predykcyjnego w zakresie prognozowania wybranych plonów

- Poprawność działania sieci typu *RBF* nauczonych na historycznych danych treningowych sugeruje nieliniowy charakter badanego problemu
- Uzyskane statystyki regresyjne dla wygenerowanych sieci pozwalają uznać, że proces uczenia przebiegał prawidłowo, a wytworzone modele neuronowe posiadają wysoką jakość, a w szczególności dobre zdolności predykcyjne.

Literatura

- [1] Rutkowska D., Piliński M., Rutkowski L. (1997). Sieci neuronowe, algorytmy genetyczne i systemy rozmyte: Wydawnictwo Naukowe PWN, Warszawa-Łódź
- [2] Osowski S. (2000). Sieci neuronowe do przetwarzania informacji: Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa
- [3] Hertz J., Krogh A., Palmer R. G. (1993). Wstęp do teorii obliczeń neuronowych: WNT, Warszawa
- [4] Boniecki P. (2004). Sieci neuronowe typu MLP oraz RGB jako komplementarne modele aproksymacyjne w procesie predykcji plonu pszenżyta: Journal of Research and Applications in Agricultural Engineering, Poznań, (1'2004), Vol. 49(1), str. 28-33
- [5] Boniecki P., Weres J. (2003). Wykorzystanie technik neuronowych do predykcji wielkości zbiorów wybranych płodów rolnych: Journal of Research and Applications in Agricultural Engineering, 4'2003, Vol. 48, str. 56-59.